

interrupting the downloading by selecting a non-sequential record; and
downloading the a non-sequential record and records sequentially thereafter until
interrupted.

✓
Please amend claim 2 as follows:

A2
2. (Amended) A method for prequeuing of files predicted to be desired by a user,
through a telecommunications link, into a local cache, based on a list generated by a restrictive
criteria, wherein the prequeuing system is responsive to any change in the sequence of user
review, such that predicted latencies for sequential file review from any given starting point are
optimized [minimized].

✓
Please add new claims 3-14 as follows:

--3. The method according to claim 1, wherein said method is executed by a browser
application.

A3
4. The method according to claim 1, wherein said method is executed by a browser
plug-in or extension.

5. The method according to claim 1, further comprising the step of cost accounting
for downloading of each record.

6. The method according to claim 1, further comprising the steps of communicating

through a network to a server hosting the records; presenting a list of records to a user; and receiving a selection of a record from the user.

7. The method according to claim 1, further comprising the steps of accounting for downloaded record; and limiting said downloading based on a predetermined parameter.

8. The method according to claim 2, wherein predicted latencies are minimized.

9. The method according to claim 2, wherein the prequeuing is optimized based on both predicted latencies and a throughput of the telecommunications link.

10. The method according to claim 2, wherein the prequeuing is optimized based on both predicted latencies and an apparent strategy for review of records by the user.

11. The method according to claim 2, wherein the prequeuing is optimized based on both predicted latencies and a cost of the record downloads.

12. The method according to claim 2, wherein the prequeuing is optimized based on both predicted latencies and a cost of on-line time.

13. The method according to claim 2, wherein the prequeuing is optimized based on both predicted latencies and a value of the user's time.

A3
09691091-101800